

# Data Mining and Visualization

In the Regulation of Medical Devices

Zhiheng Xu, Ph.D.  
FDA/CDRH



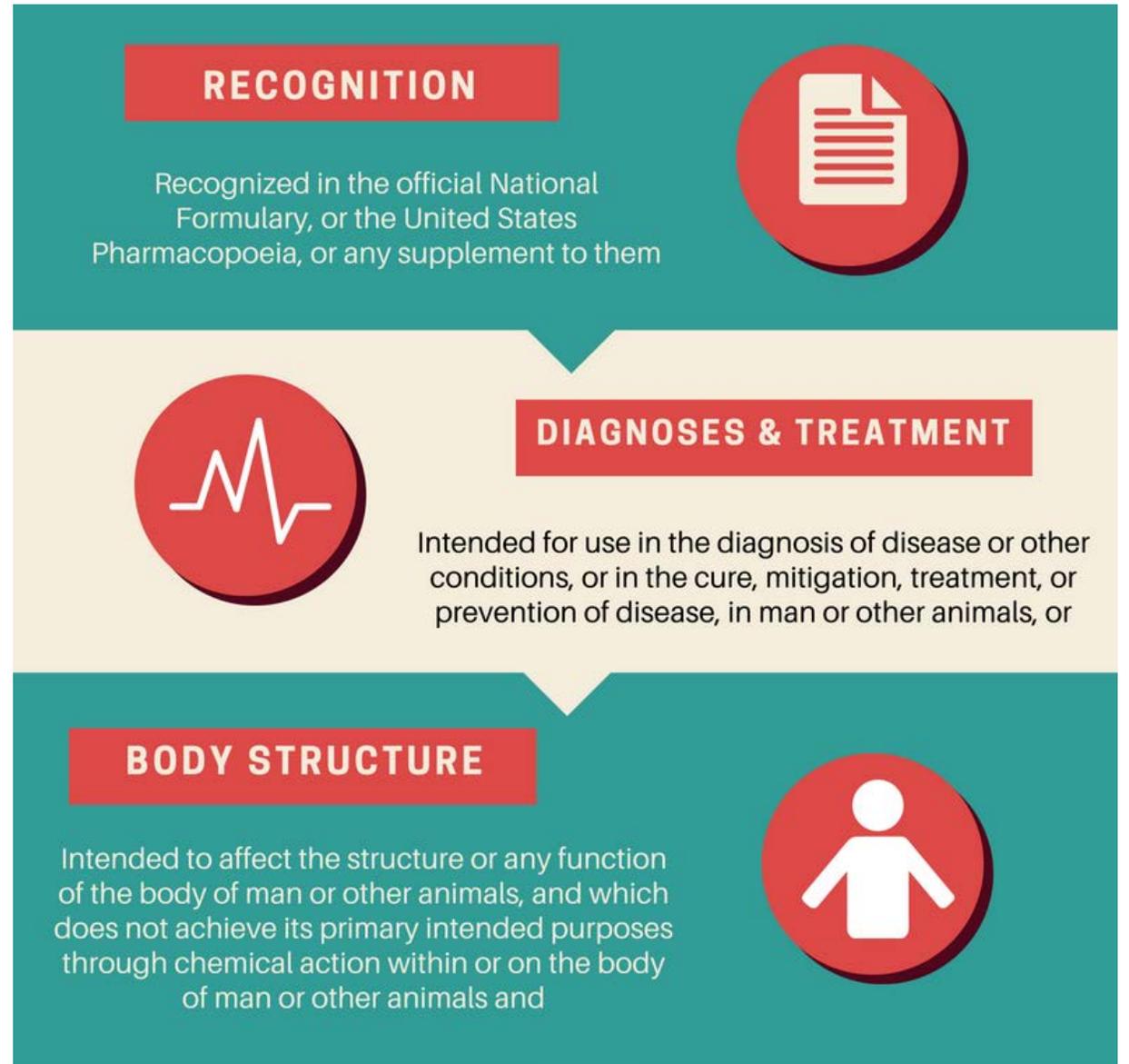
# Disclaimer

The views and opinions presented here represent those of the speaker and should not be considered to represent advice or guidance on behalf of the U.S. Food and Drug Administration.

# What is Medical Device?

“An instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article.”\*

\*: Section 201(h) of the Federal Food Drug & Cosmetic (FD&C) Act



# Device Classification



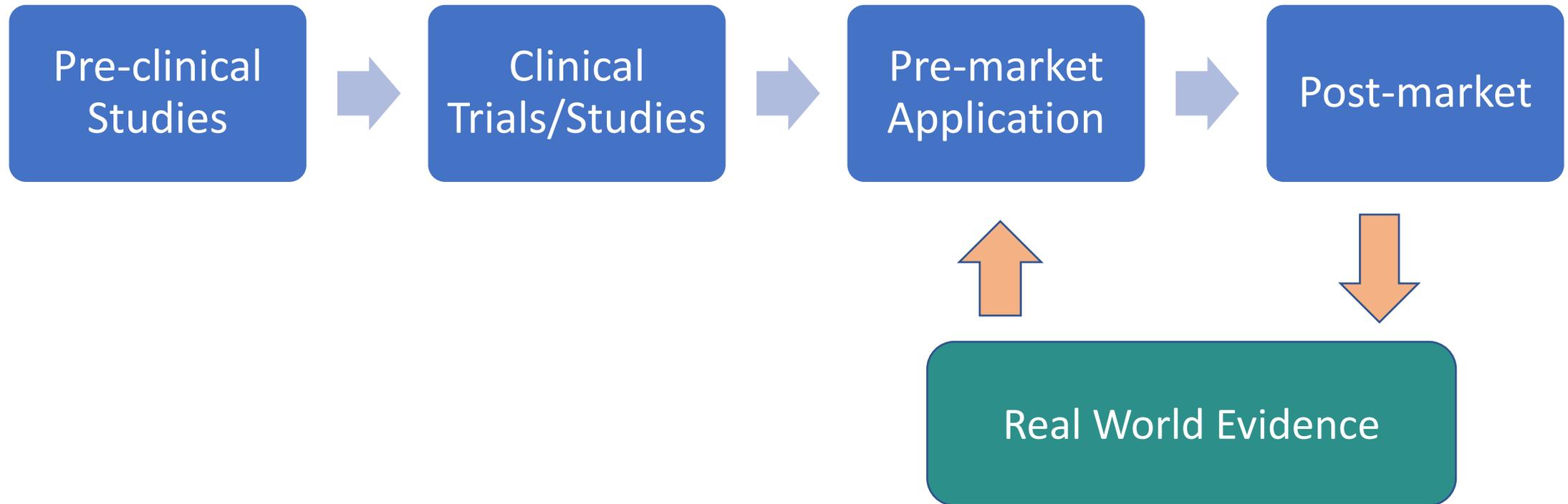
**Risk**

**Low**



**High**

# Regulatory Pathway



# Randomized Clinical Trials (RCT)



## Study Design

Protocol, Statistical Analysis Plan, Case Report Form



## Trial Implementation

Patient enrollment, data collection,



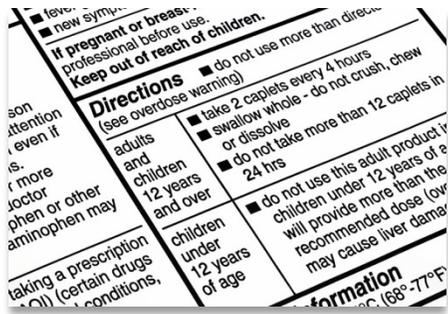
## Data Analysis

Report efficacy and safety results

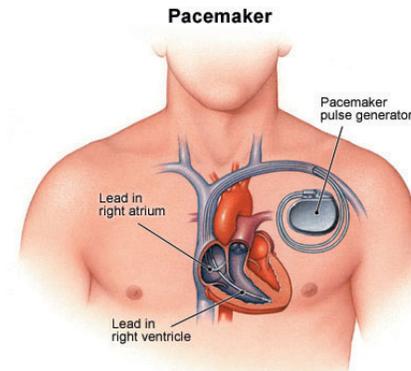
# Real World Evidence (RWE)



# Use of RWE in the Regulatory Decision-making



**Expand the labelling**



**Post-approval and post-market surveillance studies**



**Control group**



**Supplementary data**

# Benefit Risk Analysis

## Benefits

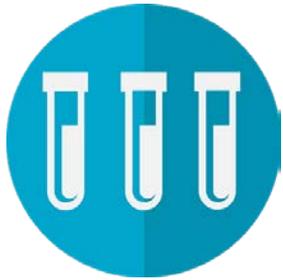
- Type of benefit(s)
- Magnitude of benefit(s)
- Likelihood of benefit
- Duration of effects
- Patient perspective on benefit
- Doctor perspective
- Medical necessity



## Risks

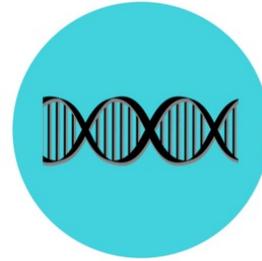
- Severity of harm
- Likelihood of risk
- False-positive or false-negative results
- Patient tolerance of risk
- Risk factors for healthcare professionals or caregivers

# Precision Medicine



## In-vitro Diagnostic (IVD)

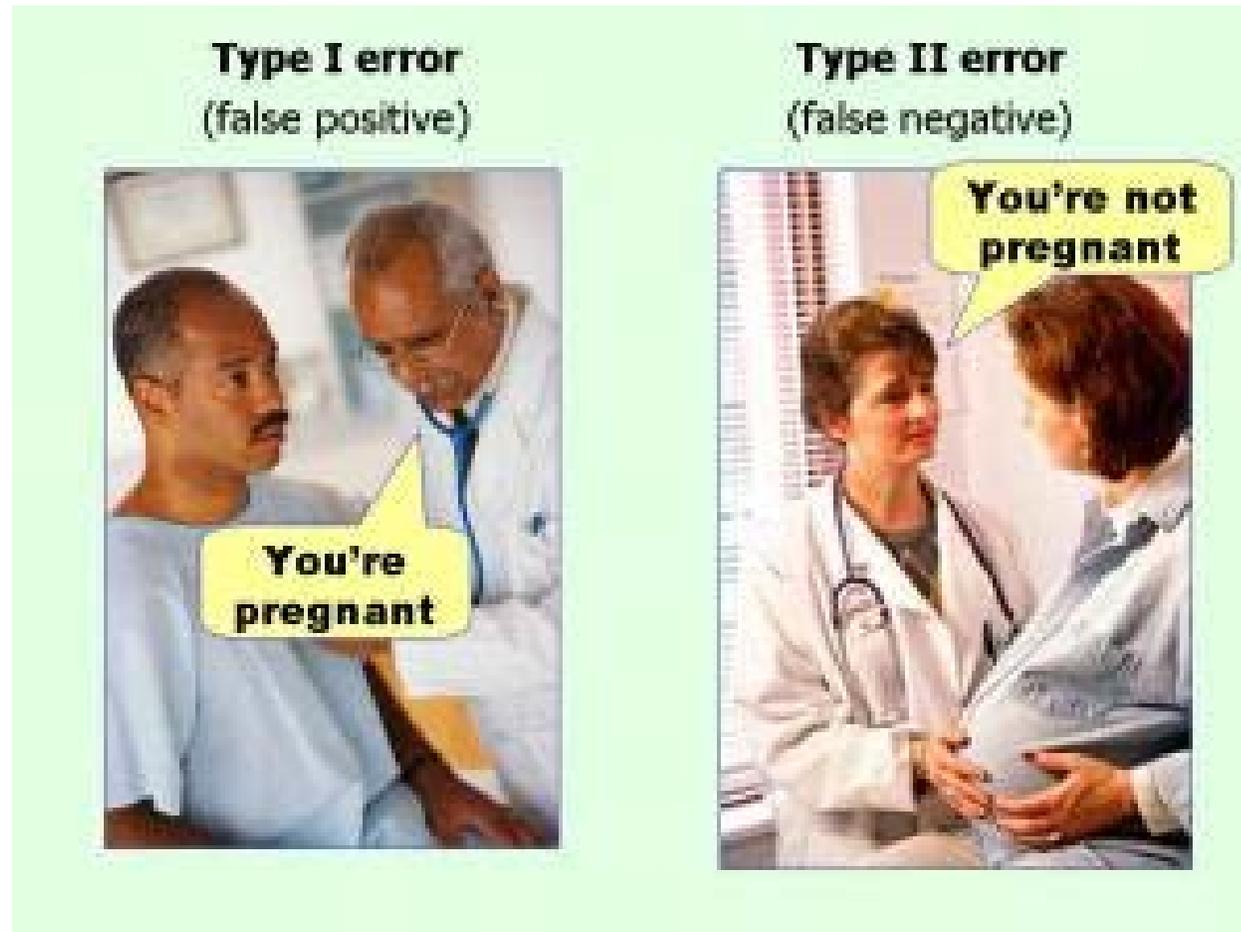
- Companion Diagnostic
- Microarray Assay
- Molecular Profiling (IHC, FISH, NGS, PCR)
- Digital Pathology



## Data Mining Challenges

- Few Data Points
- Large Dimension Spaces
- Reproducibility
- Sample Availability
- Analytical Validation
- Clinical Validation

# Type I and Type II Errors



# Post-market



Medical Device Reporting (MDR)



Mandatory Reporting



Voluntary Reporting



Manufacturer



Importer



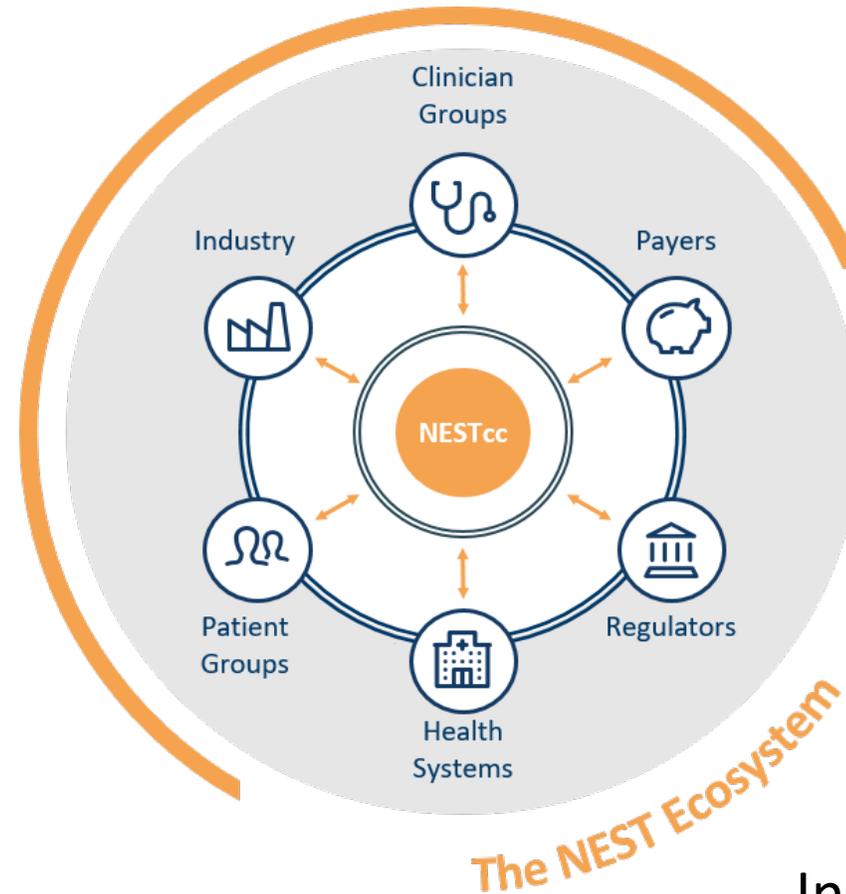
User Facility



MedWatch

# National Evaluation System for Health Technology Coordination Center (NESTcc)

[Nestcc.org](http://Nestcc.org)

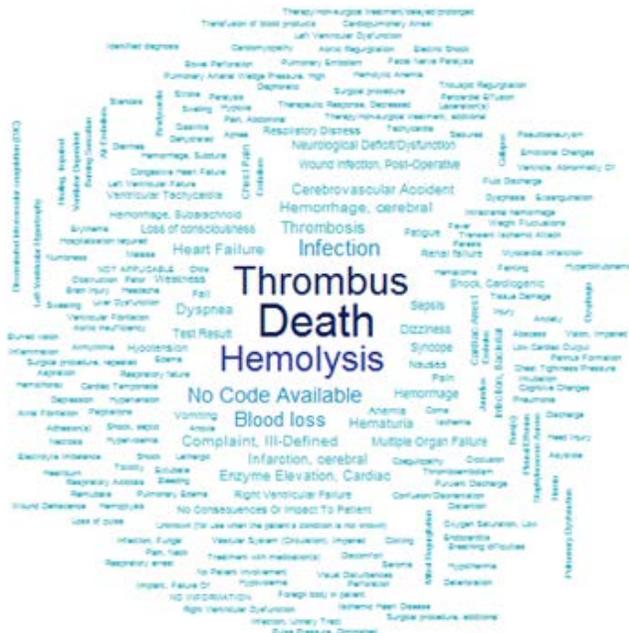


FDA/CDRH Critical Pathway

Innovative Analytical Tool for NEST  
PI: Xu, Tiwari, 2017-2019

# Signal Detection

Patient AE



Device AE



Patients' Adverse Events (AE)

Devices' Problems

# Signal Detection

AE	Device			
	1	2	...	J
1	$n_{11}$	$n_{12}$	...	$n_{1J}$
2	$n_{21}$	$n_{22}$	...	$n_{2J}$
.	.	.	...	.
.	.	.	...	.
.	.	.	...	.
I	$n_{I1}$	$n_{I2}$	...	$n_{IJ}$



An AE is a **signal** if its observed count  $\gg$  its expected count

# Signal Detection Methods

MGPS

Multi-item Gamma  
Poisson Shrinker

PRR

Proportional  
Reporting Rate

BPCNN

Bayesian confidence  
propagating neural  
network

LRT

Likelihood  
Ratio Test

# Likelihood Ratio Test (LRT)

For  $i^{\text{th}}$  AE and  $j^{\text{th}}$  device

AE	Devices		Row total
	Device j	Others	
AE i	$n_{ij}$	$n_{i,-j}$	$n_{i.}$
Others	$n_{-i,j}$	$n_{-i,-j}$	$n_{-i.}$
Col total	$n_{.j}$	$n_{.-j}$	$n_{..}$

# Likelihood Ratio Test (LRT)

$$n_{ij} \sim \text{Pois}(n_i p_i)$$

$$n_{-i,j} \sim \text{Pois}(n_{-i} \times q_i)$$

$$H_0 : p_i = q_i = p_0$$

$$H_a : p_i > q_i$$

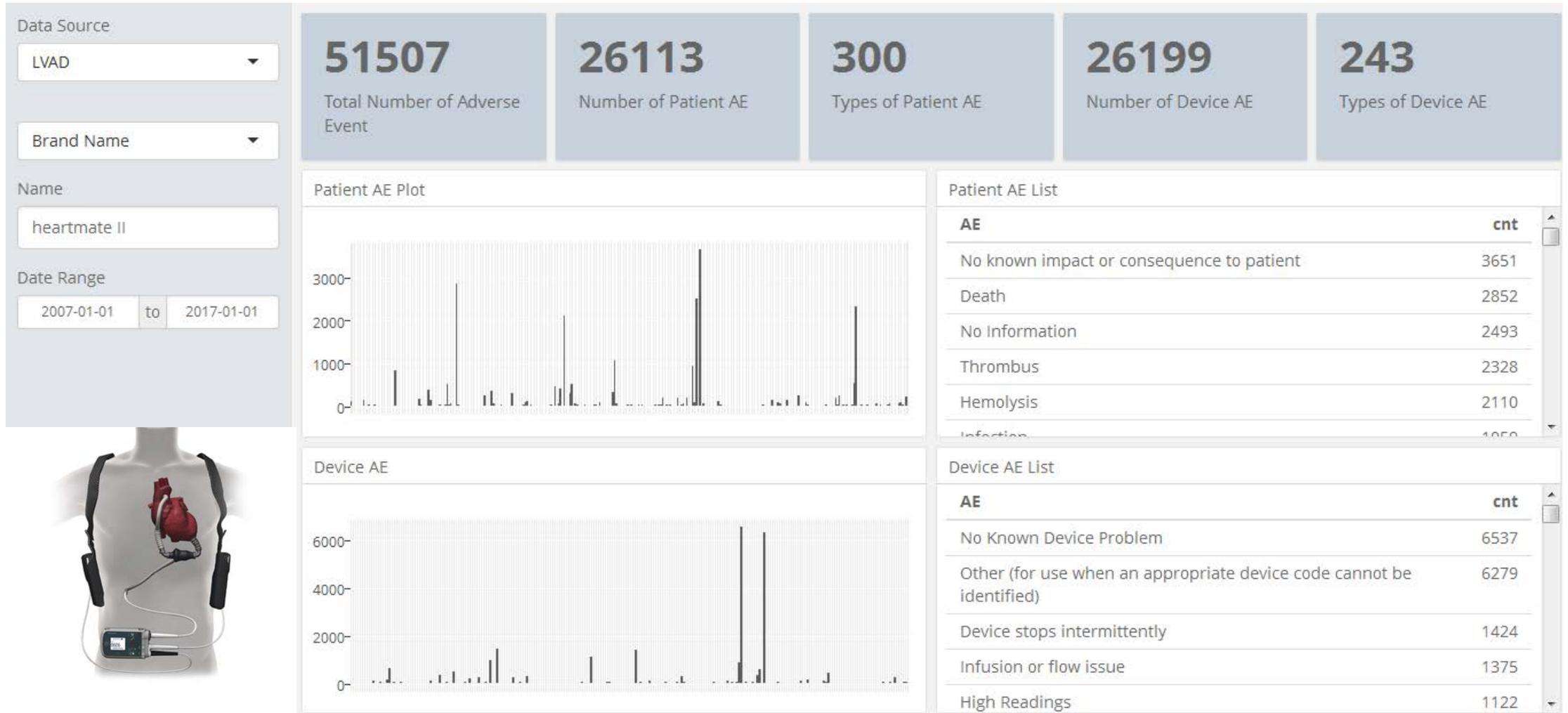
$$LR_i = \frac{L_a(\hat{p}_i, \hat{q}_i)}{L_0(\hat{p}_0)} = \frac{\left(\frac{n_{ij}}{n_i}\right)^{n_{ij}} \left(\frac{n_{-i,j}}{n_{-i}}\right)^{n_{-i,j}}}{\left(\frac{n_{\cdot j}}{n_{\cdot\cdot}}\right)^{n_{\cdot j}}}$$

- LR for every AE
- Find maximum likelihood ratio (MLR)
- Monte Carlo Simulation

# Case study: HeartMate II



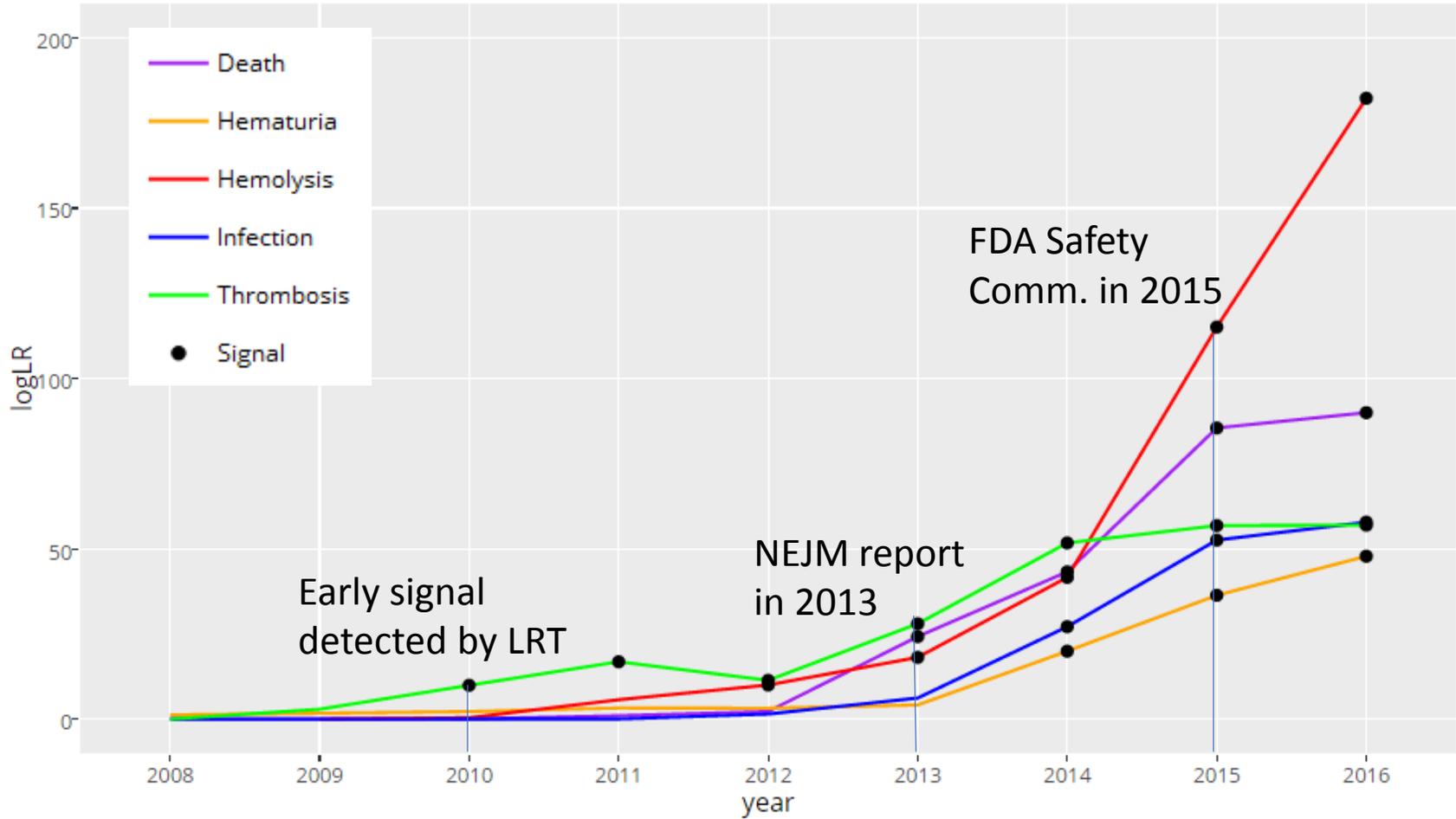
# Case study: HeartMate II



# Case study: HeartMate II

Type of AE	Reported AE Counts		logLR	pvalue
	HeartMate II only	All LVAD		
Hemolysis	2110	2684	182.32	0
Death	2852	4355	89.89	0
Complaint, Ill-Defined	508	536	81.01	0
Infection	1059	1471	57.86	0
Thrombosis	544	652	56.97	0
Enzyme Elevation, Cardiac	309	325	49.47	0
Hematuria	415	484	47.84	0
Dyspnea	344	397	41.06	0
Hemorrhage, cerebral	511	654	40.98	0
Thrombus	2328	3879	31.42	0
Infarction, cerebral	322	405	27.67	0
Heart Failure	450	615	26.32	0
Blood loss	817	1269	20.95	0
Fatigue	112	139	10.11	0.0007
Chest Pain	129	166	9.99	0.0008
Pain	118	151	9.36	0.0014
Shock, Cardiogenic	76	89	8.55	0.0048
Right Ventricular Failure	141	197	7.25	0.0204
Hemorrhage	288	455	6.27	0.0491

# Safety Signals Over Time – Cumulative Year Total





# Bioresearch Monitoring (BIMO) Site Selection

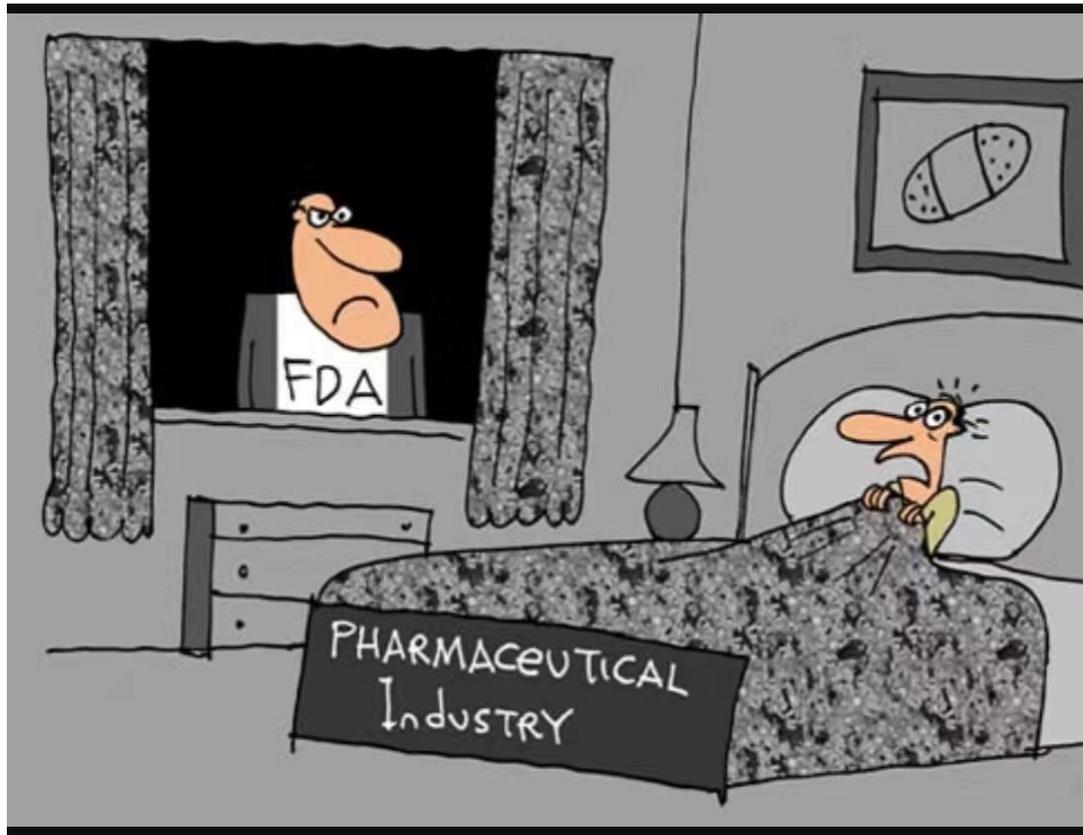


Image source: Pharmaceutical Manufacturing

# BIMO Site Selection



Multi-Center Clinical Trials

# BIMO Site Selection – A hypothetical example

Site ID	Sample Size	Serious AE event	Protocol deviation event	Death event	Success rate	Non-Serious AE
1	4	1	2	0	2	4
2	24	15	21	0	18	24
3	3	1	0	1	2	3
4	8	4	7	0	5	8
...	...	...	...	....	...	...
17	16	13	12	1	9	15
18	2	2	0	0	2	2
19	1	0	1	0	1	1
20	8	2	5	0	7	2

# BIMO Site Selection

Statistical test for each variable and p-value generation from statistical tests by variable

Serious AE				Non-Serious AE				Death				Protocol Deviation				Success Rate			
Site	Size	...	P-value	Site	Size	...	P-value	Site	Size	...	P-value	Site	Size	...	P-value	Site	Size	...	P-value
1	4		0.6	1	4		0.8	1	4		0.9	1	4		0.5	1	4		0.5
2	24		<b>0.02</b>	2	24		<b>0.03</b>	2	24		<b>0.01</b>	2	24		<b>0.04</b>	2	24		<b>0.02</b>
3	3		0.9	3	3		0.7	3	3		0.9	3	3		0.3	3	3		0.3

Combination of p-values in a p-value Matrix

P-value matrix

Site	P-value	P-value	P-value	....	P-value	P-value
1	0.6	0.8	0.9		0.5	0.5
2	<b>0.02</b>	<b>0.03</b>	<b>0.01</b>		<b>0.04</b>	<b>0.02</b>
3	0.9	0.7	0.9		0.3	0.3

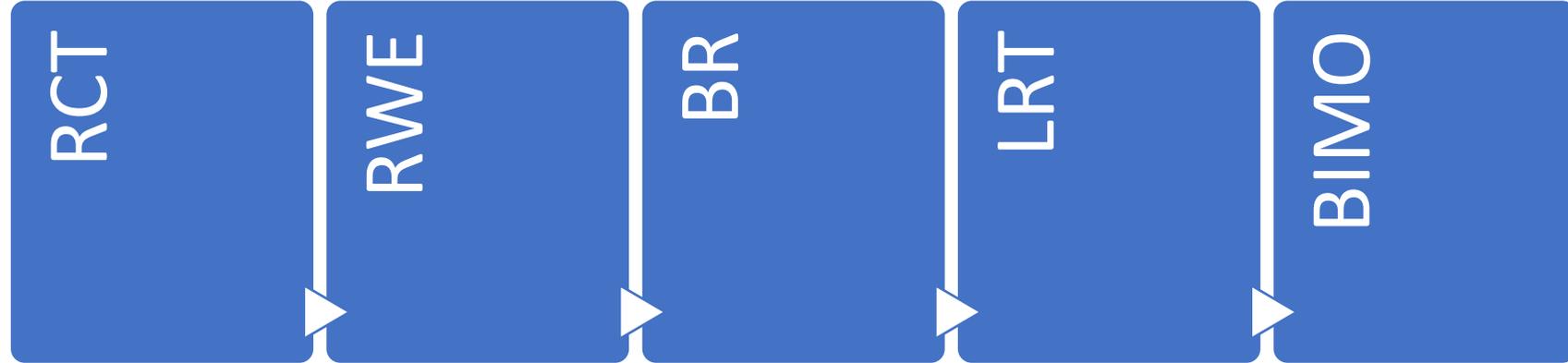
Site ranking

Site	Size	...	P-value	Ranking
1	4		0.5	3
2	24		<b>0.01</b>	<b>1</b>
3	3		0.1	2



Image Source: [www.youtube.com/watch?v=nIsLFgsC1qY](https://www.youtube.com/watch?v=nIsLFgsC1qY)

# Conclusion



Thank you!

# Acknowledgement

The author would like to thank Drs. Ram Tiwari, Lan Huang, Howard Yao, Jianjin Xu, Meijuan Li and Changhong Song and others from CDRH/Division of Biostatistics for their inputs.